复杂大田场景下基于改进YOLOv8的小麦幼苗期叶片数快速检测方法

侯依廷[1,2,3]，饶　元[1,2,3*]，宋　贺[4]，聂振君[1,2,3]，王　坦[1,2,3]，何豪旭[1,2,3]

（1. 安徽农业大学 信息与人工智能学院，安徽合肥 230036，中国；2. 农业农村部农业传感器重点实验室，安徽合肥 230036，中国；3. 智慧农业技术与装备安徽省重点实验室，安徽合肥 230036，中国；4. 安徽农业大学 农学院，安徽合肥 230036，中国）

摘　要：［**目的/意义**］小麦叶片数是衡量植株生长状况、确定茎蘖动态、调节群体结构的重要指标之一。目前大田环境下小麦叶片计数主要依靠人工、耗时耗力，而现有的自动化检测计数方法的效率与精度难以满足实际应用需求。为提高小麦叶片数检测的准确性，设计了一种复杂大田环境下高效识别小麦叶尖的算法。［**方法**］本研究以手机和田间摄像头获取的可见光图像构建了两种典型光照条件下出苗期、分蘖期、越冬期等多个生长期的小麦叶片图像数据集。以YOLOv8为基础网络，融合坐标注意力机制降低背景环境的干扰，提高模型对小麦叶尖轮廓信息的提取能力；替换损失函数加快模型收敛速度；增加小目标检测层提高对小麦叶尖的识别效果，降低漏检率。设计了一种适用于叶尖小目标识别的深度学习网络，通过检测图像叶尖数量从而得出叶片数。［**结果与讨论**］本研究提出的方法对小麦叶尖的识别精确率和$mAP_{0.5}$分别达到91.6%和85.1%，具有良好的检测效果。在复杂大田环境下该方法具有更好的适应能力，能够在不同光照条件下实现自适应检测，模型鲁棒性强。小麦幼苗期叶片检测漏检率低，说明该方法能够满足复杂大田场景下小麦叶尖识别的需求，提高了小麦叶片数检测的准确性。［**结论**］本研究可为复杂大田场景下小麦叶片数检测的研究提供参考，为小麦长势高质量评估提供技术支撑。

关键词：小麦叶片；叶尖识别；叶片计数；注意力机制；YOLOv8；深度学习

中图分类号：S24　　**文献标志码：**A　　**文章编号：**SA202403019

引用格式：侯依廷，饶元，宋贺，聂振君，王坦，何豪旭．复杂大田场景下基于改进YOLOv8的小麦幼苗期叶片数快速检测方法[J]. 智慧农业(中英文)，2024，6(4)：128-137. DOI：10.12133/j.smartag.SA202403019

HOU Yiting, RAO Yuan, SONG He, NIE Zhenjun, WANG Tan, HE Haoxu. A Rapid Detection Method for Wheat Seedling Leaf Number in Complex Field Scenarios Based on Improved YOLOv8[J]. Smart Agriculture, 2024, 6(4): 128-137. DOI：10.12133/j.smartag.SA202403019 (in Chinese with English abstract)

0　引　言

随着现代农业技术的快速发展，作物生长状态的实时监测和精准管理成为了提高作物产量和质量的关键[1]。小麦作为全球重要的粮食作物之一，其生长监测、产量预测以及田间管理优化具有重要意义[2]。小麦生长过程中的叶片数是衡量植株生长状况、预测产量的重要指标之一[3]。小麦幼苗期的叶片数量是决定植株养分吸收效率及光合作用能力的关键因素，同时对小麦分蘖的形成以及最终产量的提高起着决定性作用，是确定小麦茎蘖结构、调节群体结构的重要指标。纵观小麦叶片检测方法的发展历程，可以得出：基于人工计数的方法具有主观性强、效率低且劳动强度大的弊端；基于仪器的检测方法获得结果受外界环境等因素影响；基于数理统计的检测方法工作量大，受限于专业的统计学和农学知识。

近年来，随着计算机算力的不断提升，深度学习技术在目标检测领域广泛应用[4,5]。与传统的目

收稿日期：2024-03-15

基金项目：国家自然科学基金项目（32371993）；安徽省重点研究与开发计划项目（202204c06020026，2023n06020057）；安徽省高校自然科学研究重大项目（2022AH040125，2023AH040135）

作者简介：侯依廷，研究方向为农业信息化。E-mail：wjx0626fmvp@163.com

*通信作者：饶　元，博士，教授，研究方向为农业信息化。E-mail：raoyuan@ahau.edu.cn

标检测算法相比，深度卷积神经网络（Convolutional Neural Network, CNN）具有强大的特征学习能力，能够从海量的训练数据中自主挖掘多层次的特征表达[6,7]。基于卷积神经网络的目标检测算法如YOLO（You Only Look Once）和SSD（Single Shot Detector）一阶检测器系列和Faster R-CNN（Faster Region-Based Convolutional Neural Network）两阶检测器系列已被证实在对作物的识别上取得了不错的效果[8-10]。众多学者不断优化和改进这些深度学习算法模型，并尝试将其应用于叶片识别与计数的研究中。Xu等[11]提出了一种结合半监督学习、深度学习以及无人机数字成像技术用于田间玉米幼苗叶片计数的方法，对于完全展开的叶片和新出现的叶片，计数准确率分别为69.4%和72.9%。Deb等[12]提出名为LC-Net的新型CNN模型，通过结合叶片局部特征和全局图像信息，在公开测试集上的叶片计数精确率优于现有技术。Praveen和Domnic[13]提出了一种提取植物数字图像中叶片区域并计算叶片数量的有效方法，在叶子分割挑战的基准数据集上实现了95.4%的分割精度。Yigit等[14]利用神经网络和支持向量机（Support Vector Machine, SVM）设计了一种植物叶片的自动识别器，SVM模型识别准确率为92.91%。Xie等[15]基于一种CNN提出深度回归模型，降低训练成本的同时提高了叶片计数的准确性。Kolhar和Jagtap[16]使用U-Net和分水岭算法实现了植物叶片的分割与计数，平均叶片计数误差为0.26。

国内研究者运用深度学习方法针对叶片识别与计数的研究取得了一定成效，但依旧存在一些问题。例如，对小麦叶片数检测研究较少，大田背景复杂会使算法将背景误判，缺乏对于特定场景下的小麦叶片数检测研究。因此，本研究利用手机和田间摄像头获取的可见光图像构建数据集，对小麦叶尖进行标注，提出基于YOLOv8-CSD（You Only Look Once-CA、SIoU、Small Object Detection Layer）模型的小麦叶片数检测方法。通过替换损失函数加快模型收敛速度，嵌入坐标注意力机制（Coordinate Attention, CA）强化小麦叶尖目标特征，增加小目标检测层以拓宽模型的感受野，提高小麦叶尖识别准确性，从而通过检测叶尖数量推断出小麦叶片数。该方法为小麦叶片数快速检测提供了可行方案，为复杂大田场景下小麦的生长监测提供参考依据，以便于更精确地评估小麦的生长情况。

1　试验材料

1.1 图像采集

小麦数据集采集地点位于安徽省合肥市庐阳区合肥高新技术农业园，采集时间为2023年11月19日至2024年2月10日。使用手机OPPO Reno10基于小麦分蘖节25、35和50 cm高度处分别拍摄图像。采用海康威视DS-2DE4423IW-D/GLT/XM在放大5倍、10倍、15倍、20倍情况下分别采集图像，设备距地面高度为3.5 m。图1展示了小麦田间场景，以及手机、田间摄像头拍摄的田间小麦图像。



a. 小麦田间场景



b. 手机拍摄的越冬期小麦图像



c. 田间摄像头拍摄的分蘖期小麦图像

图1　不同设备拍摄的小麦图像数据示例

Fig.1 Example of wheat images captured by different devices

为实现小麦叶尖识别与叶片数检测的目标，分别在小麦出苗期、分蘖期、越冬期等幼苗关键生长阶段采集图像，全面捕捉小麦幼苗期不同生长阶段的形态特征。复杂大田场景下采集的图像涵盖多样的背景元素、小麦的生长变化，以及不同的光照条件。如图1b所示，由于杂草以及残留作物秸秆共同构成了复杂的视觉环境，难以分割，严重干扰了小麦叶片的准确检测。因此，对采集的小麦图像通过调整对比度等操作进行数据增强处理。鉴于大田小麦采用撒播方式种植，密度大，在分蘖期以后会

出现叶片密集相互遮挡的问题，获取叶片图像数据时，手机平行于地面拍摄，摄像头采用垂直俯视角度拍摄。由于自然光照条件的不断变化，如日照角度和强度的波动以及由此产生的阴影效果，进一步增加了图像处理的难度，因此，选择在晴天和阴天分别采集数据，提升模型在实际应用中的泛化性。表1为不同生长期不同光照条件下获取的小麦图像数据集。

表1 不同设备拍摄的不同生长期和光照条件的小麦图像数据组成

Table 1 Composition of wheat image data captured by different devices with different growth stages and lighting conditions

数据来源	不同生长期小麦图像数量/张			不同光照条件小麦图像数量/张	
	出苗期	分蘖期	越冬期	晴天	阴天
手机	187	323	254	354	410
田间摄像头	—	62	44	48	58
总数	187	385	298	402	468

注："—"表示无。

1.2 数据集构建

综合考虑硬件条件和试验参数的需要，同时为了便于模型训练，利用Python语言编写的随机裁剪缩放函数将手机获取的原始图像裁剪成分辨率像素640×640图像，将田间摄像头获取的分辨率像素为2 560×1 440的可见光图像缩放至像素为640×640的输入模型。小麦叶尖形态特征明显，相对容易识别，故利用LabelImg标注软件对进行预处理过后的小麦叶片图像进行边界框标注，通过最小外接矩形标注出小麦叶尖。

为增强数据集的多样性，提高模型泛化能力，实现小麦叶尖识别和叶片数检测的目标，本研究选取手机原始图像数据764张，经旋转后，每张原始图像被扩充成2张像素为640×640的图像，共计1 528张。考虑到图像质量不高会造成特征提取困难，手动剔除了过度曝光、质量较低的图像。经过筛选后，剩余1 426张图像构成数据集。选取田间摄像头获取的106张原始图像，经过缩放裁剪后共计420张放入数据集。将数据集按照8：2的比例划分为训练集和剩余集，在剩余集中进一步平均随机划分，将其中的一部分作为验证集，另一部分作为测试集，形成最终数据集。其中，训练集1 477张，验证集185张，测试集184张，共标记出31 748个小麦叶尖。

2 算法模型构建和评价指标

2.1 YOLOv8算法

YOLOv8是Ultralytics公司开源的一种单阶段目标检测算法，该系列模型有n、s、m、l和x这5个版本，模型参数量依次增加。其主干网络[17]由卷积模块（ConvModule, CONV）、C2f（CSPLayer_2Conv）模块、空间金字塔池化（Spatial Pyramid Pooling -Fast, SPPF）模块组成。CONV模块封装了三个功能，包括卷积（Conv2d）、批量归一化层（Batch Normalization, BN）和激活函数SiLU[18]，其作用是对输入的特征图执行卷积操作，从而提取特征，识别目标。C2f模块用于将高层特征与低层特征进行融合，以获取更丰富的语义信息。SPPF模块旨在通过将不同尺度的特征汇集到固定大小的特征图中来加快网络的计算。Neck网络保留使用路径聚合网络-特征金字塔网络（Feature Pyramid Networks- Path Aggregation Network, PAN-FPN）结构，加强了特征信息。Head网络采用解耦头结构，将分类和检测头分离；采用无锚分体式检测头，保证在较短时间对图像物体进行识别和检测，有效地提高了检测速度。

2.2 嵌入CA注意力机制

注意力机制通过让深度学习网络更加关注需要关注的地方，提升神经网络的性能，目前广泛应用于计算机视觉任务中[19]。但常规注意力机制易忽略目标空间位置信息[20]。由于复杂大田场景下杂草难以分割，小麦叶片与杂草颜色相似性高，导致识别准确性下降。为提高模型对小麦叶尖轮廓信息的提取能力，减少背景环境的干扰，在YOLOv8m模型中嵌入CA注意力机制。该机制充分考虑了通道信息和方向相关的位置信息的关联性，能够使感兴趣的区域更准确地被捕获，提升模型特征提取能力，帮助网络更好地识别目标[21]。

CA注意力机制对输入特征图进行两个一维全局池化操作，分别得到宽度和高度方向上的特征映射，接着将两个特征映射合并形成一个新的特征层。对合并后的特征层进行卷积操作和激活操作进一步获得特征图。之后沿着空间维度将新获得的特征图进行split操作分离出宽度和高度方向的特征层，结合Sigmoid激活函数[22]得到在宽度与高度方向上的注意力向量。最后，将原始输入特征图进行加权计算，输出在宽度和高度方向上带有注意力权

重的特征图。

2.3 优化损失函数

YOLOv8算法中，预测框回归损失采用完全交并比（Complete Intersection over Union, CIoU）损失函数[23]，该函数主要关注了检测框的尺度损失，而没有考虑到真实框与预测框之间的方向，导致模型收敛速度较慢，影响整体检测性能。由于复杂大田场景下分蘖期以后的小麦叶片密集相互遮挡，叶尖识别变得困难。为了加快预测框回归到真实框的速度，提高对小麦叶尖的识别效果，本研究采用形状交并比（Shape Intersection Over Union, SIoU）损失函数[24]替换CIoU损失函数。公式（1）为SIoU损失函数，其包含公式（2）角度损失Λ、公式（3）距离损失Δ、公式（4）形状损失Ω。图2为SIoU参数示意图。

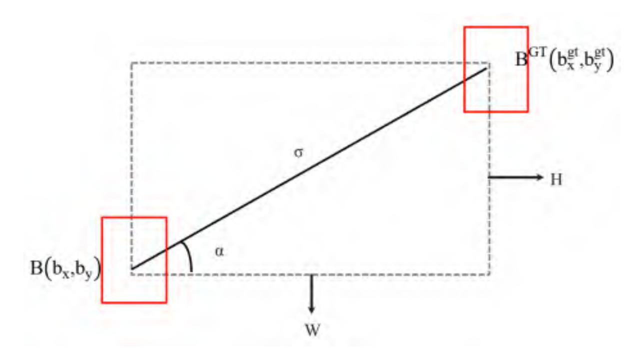
$$Loss_{SIoU} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \tag{1}$$

$$\Lambda = \cos\left(2\left(\arcsin\left(\frac{H}{\sigma}\right)\right) - \frac{\pi}{4}\right) \tag{2}$$

$$\Delta = 2 - e^{-\gamma p_x} - e^{-\gamma p_y} \tag{3}$$

$$\Omega = \sum_{t = w,h}(1 - e^{-W_t})^{\theta} \tag{4}$$

式中：IoU（Intersection over Union）为预测框与真实框之间的交并比；H为预测框和真实框中心点的横坐标之差；σ为真实框和预测框中心点距离，px；γ为赋予时间的距离值；ρ_x和ρ_y分别为真实框与预测框中心点最小外接框宽比值的平方和最小外接框高比值的平方；θ为控制对形状损失的关注程度，参数范围为［2，6］；W为预测框和真实框中心点的纵坐标之差。



注：B和B^{GT}点分别为预测框和真实框中心点；H和W分别为B和B^{GT}点横坐标与纵坐标之差；σ为真实框和预测框中心点距离；α为两中心点连线与水平线的夹角。

图2　SIoU参数示意图

Fig.2　SIoU parameters schematic diagram

2.4 添加小目标检测层

原始的YOLOv8模型使用三种不同尺寸的特征图进行目标检测：80×80像素的特征图用于检测像素大小在8×8以上的目标；40×40像素的特征图用于检测像素大小16×16以上的目标；20×20像素的特征图用于检测像素大小在32×32以上的较大目标。然而，由于出苗期小麦叶尖尺寸较小，其宽度和高度均小于8像素，原始模型可能无法准确捕捉到网格内的小麦叶尖特征信息。因此，通过增加小目标检测层，即在原始模型中新增一个像素大小为160×160的检测特征图，充分利用底层特征的高分辨率信息和深层特征的高语义信息，在保持计算效率的同时，让网络更加关注小麦叶尖的特征信息，从而提高小麦叶尖的检测效果。

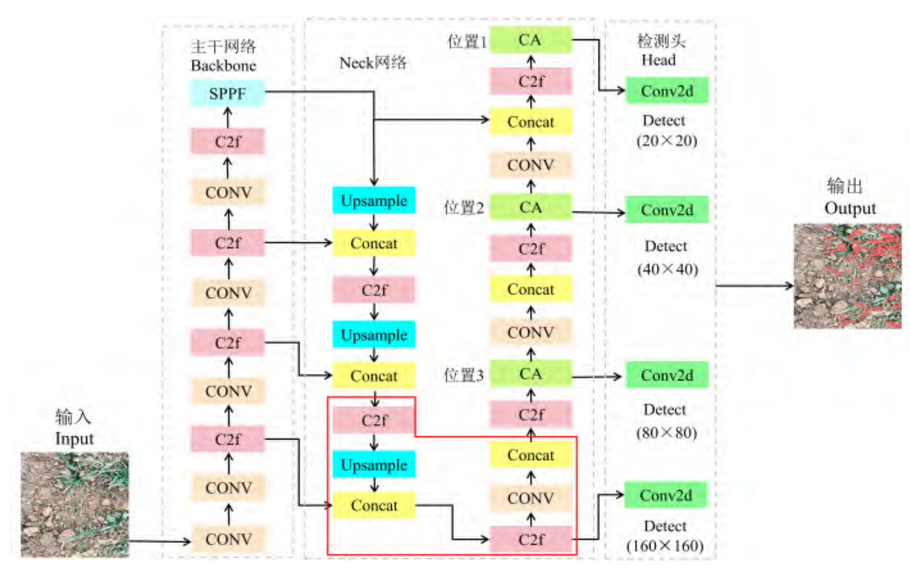
2.5 小麦叶尖识别算法模型

考虑到实际生产条件下平台算力的限制，为了实现小麦叶尖识别精度与检测速度的平衡，本研究基于YOLOv8m网络进行改进，提出了YOLOv8-CSD算法模型。YOLOv8-CSD中嵌入CA注意力机制来突出重要的通道特征，增强小麦叶尖识别能力；替换损失函数加快模型训练收敛速度；添加小目标检测层提升模型对细节的敏感性，解决较小叶尖的识别问题。YOLOv8-CSD检测头包括4个检测层，具体如表2所示。其中，输出特征图像素为20×20的检测层具有较大的感受野，适合识别大尺寸目标。输出特征图像素为160×160的检测层具有较小的感受野，适合捕捉小尺寸目标。40×40像素和80×80像素的检测层，分别针对检测中等尺寸和较小尺寸的目标。YOLOv8-CSD网络结构如图3所示。

表2　YOLOv8-CSD模型检测头组成

Table 2　Composition of detection head for YOLOv8-CSD model

特征图/像素	20×20	40×40	80×80	160×160
感受野	大	中	较小	小
锚框	（116，90）（156，198）（373，326）	（30，61）（62，45）（59，119）	（10，13）（16，30）（33，23）	（5，7）（9，13）（11，15）



注:红色实线框为小目标检测层添加位置;位置1、位置2、位置3为添加CA注意力机制位置。

图3 YOLOv8-CSD网络结构图

Fig.3 YOLOv8-CSD network structure diagram

2.6 模型训练与评价指标

2.6.1 模型训练环境

本研究使用Windows 10操作系统,内存为32 G,显卡型号为NVIDIA GeForce RTX 4060 8 GB,搭载13th Gen Intel(R) Core(TM) i5-13500HX处理器,深度学习框架采用PyTorch1.12,编程平台为PyCharm,编程语言为Python3.10,所有对比算法均在相同环境下运行。

实验图像输入像素尺寸640×640,并配合使用Mosaic数据增强。在模型训练过程中,初始学习率为0.01,配合优化器Adam(Adaptive Moment Estimation Optimizer)调整学习率。设定权重衰减率为0.000 5,动量因子为0.937。在进行性能测试时,预测的结果采用非极大值抑制。

2.6.2 评价指标

本研究采用精确率(Precision, P)、召回率(Recall, R)、$mAP_{0.5}$(Mean Average Precision)以及模型大小等指标进行模型评价,各评价指标计算如公式(5)和公式(6)所示。P表示在所有模型预测为小麦叶尖的结果中,预测正确数量占总数的比例;R表示模型预测为小麦叶尖的数量占实际叶尖数量的比例。

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{5}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

式中:TP、FP、FN分别为真正例、假正例、假反例数量;以R为横轴,P为纵轴,PR曲线下的面积定义为平均精度(Average Precision, AP),AP越大,说明检测结果越好。$mAP_{0.5}$是AP值在所有类别下的均值。

3 结果与分析

3.1 消融实验

在消融实验中,为保证对比合理性,所有模型均使用相同的数据集,并在相同的参数条件下进行训练。为对比使用CA注意力机制模块、替换损失函数模块和增加小目标检测层模块对原YOLOv8模型带来的性能提升表现,进行模型消融实验,对比结果如表3所示,可知在YOLOv8中添加CA注意力机制,与原模型相比,召回率和$mAP_{0.5}$分别提高0.9%和1.4%;替换损失函数SIoU后,召回率和$mAP_{0.5}$分别提高了0.5%和0.4%;通过添加小目标检测层,与原YOLOv8模型对比,召回率和$mAP_{0.5}$分别提高了3.2%和3.1%;改进的YOLOv8模型对小麦叶尖的识别精确率、召回率、$mAP_{0.5}$分别为91.6%、80.0%和85.1%。与原YOLOv8相比,分别提高2.3%、6.4%和3.3%。

3.2 不同网络模型识别性能对比实验

为验证本研究提出的YOLOv8-CSD算法模型在复杂大田场景下小麦叶片数据集上检测的有效性,选择当前主流的二阶网络模型Faster R-CNN、Mask R-CNN (Mask Region-based Convolutional

表3　小麦叶尖识别研究中改进YOLOv8模型的消融实验

Table 3 Improved YOLOv8 model ablation experiment for wheat leaf tip recognition research

序号	添加CA注意力机制	替换SIoU损失函数	添加小目标检测层	P/%	R/%	$mAP_{0.5}$/%
1	×	×	×	89.3	73.6	81.8
2	√	×	×	90.2	74.5	83.2
3	×	√	×	89.7	74.1	82.2
4	×	×	√	90.0	76.8	84.9
5	√	√	√	91.6	80.0	85.1

注：√表示使用该模块；×表示没有使用该模块。

Neural Network）[25]与一阶网络模型YOLOv7[26]、YOLOv8、SSD共6种网络模型在相同的训练环境下进行对比实验。选择评价指标从精确率、召回率、$mAP_{0.5}$得到对比实验结果，如表4所示，所有数据均为平均值。从表4中得出，二阶卷积神经网络Faster R-CNN、Mask R-CNN检测效果较差，$mAP_{0.5}$远低于其他一阶网络模型，可能在特征提取阶段不如一阶模型优化得当，导致在复杂背景的小目标检测上的性能较差；本研究提出的YOLOv8-CSD的$mAP_{0.5}$能够达到85.1%，与原YOLOv8相比，提升3.3%，召回率提高6.4%；与YOLOv7和SSD相比，$mAP_{0.5}$分别提高了13.3%和13.6%。YOLOv8-CSD模型在小麦叶尖识别精度方面优于其他模型，模型占用内存较小且召回率大幅度提高，说明改进后的模型能够减少背景因素的干扰，对小麦叶尖的定位能力更好，因此本研究提出的模型更适用于复杂大田场景下小麦叶尖识别与叶片数的检测。

不同模型在数据集上的可视化实验结果如图4所示，可看出，本研究提出的模型收敛速度较快，并且能够在短时间内训练出较优的模型。

3.3 不同设备获取图像检测效果对比

为探究YOLOv8-CSD模型对不同设备获取的

表4　小麦叶尖识别研究中不同模型的检测性能对比平均值

Table 4 Comparison of detection performance average of different models in wheat leaf tip recognition research

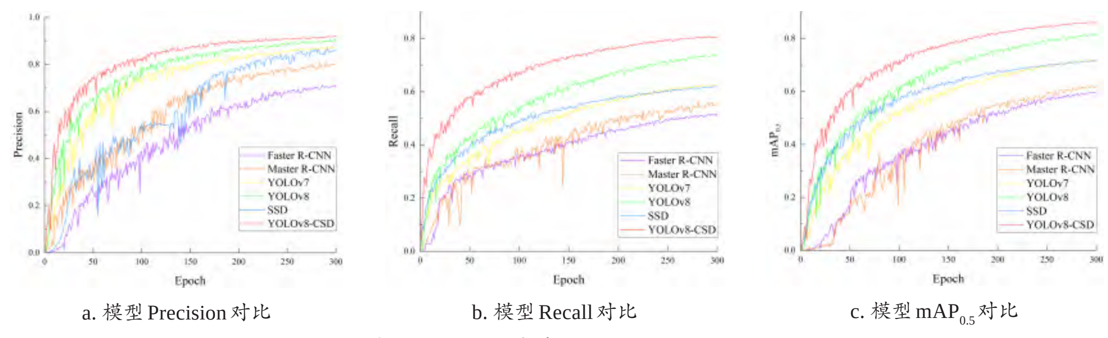
模型	P/%	R/%	$mAP_{0.5}$/%	模型大小/MB
Faster R-CNN	70.9	50.4	58.7	108.2
Mask R-CNN	79.3	54.8	61.6	244.0
YOLOv7	87.2	62.5	71.8	74.8
YOLOv8	89.3	73.6	81.8	49.7
SSD	85.6	61.9	71.5	78.1
YOLOv8-CSD	91.6	80.0	85.1	52.2

小麦叶片图像的检测效果，对验证集中143张手机拍摄图像和42张田间摄像头拍摄图像进行对比试验。选择识别精确率、召回率、$mAP_{0.5}$作为评价指标，得到YOLOv8-CSD模型对不同设备获取图像的具体检测结果。根据表5可知，模型在检测手机拍摄的小麦叶片图像时，叶尖识别的$mAP_{0.5}$达到了86.78%，而田间摄像头拍摄图像叶尖识别的$mAP_{0.5}$为82.43%。田间摄像头拍摄图像的叶尖识别精度与手机相比偏低，这是由于两种设备获取的图像空间分辨率大小不一样、采用的预处理方式不一样以及摄像头平行地面拍摄范围有限，导致处理后图像数据所包含的小麦叶尖数量有一定的差距，因此在进行小麦叶尖识别时会出现不同设备的识别精度差。但YOLOv8-CSD模型在检测两种不同设备获取的小麦叶片图像时，叶尖识别的$mAP_{0.5}$均超过80%，说明应用YOLOv8-CSD模型能够有效识别两种不同设备拍摄的小麦叶尖。

表5　YOLOv8-CSD模型对不同设备获取的小麦叶片图像检测结果对比

Table 5 Comparison of detection results of wheat leaf images obtained from different devices using YOLOv8-CSD model

图像采集设备	P/%	R/%	$mAP_{0.5}$/%
田间摄像头	90.18	78.16	82.43
手机	92.43	83.51	86.78



a. 模型Precision对比　　b. 模型Recall对比　　c. 模型$mAP_{0.5}$对比

图4　小麦叶尖识别研究中不同模型检测性能对比图

Fig. 4 Comparison of detection performance of different models in wheat leaf tip recognition research

3.4 模型改进前后复杂环境检测效果对比

本研究旨在探究YOLOv8-CSD模型在复杂大田环境下对小麦叶尖的检测效果。考虑到手机能够从多角度捕捉包含杂草干扰及叶片相互遮挡情况的小麦图像，本研究从测试集中随机选取手机拍摄的小麦叶片图像，分别输入YOLOv8和YOLOv8-CSD模型进行对比实验。检测结果如图5所示。在杂草及叶片相互遮挡的复杂条件下，YOLOv8-CSD模型较YOLOv8模型检测置信度更高，对于小麦叶尖的识别更准确。在图5中，存在叶片遮挡情况的原始图像中标注了被遮挡区域，YOLOv8未能识别出，而YOLOv8-CSD能够有效识别，说明本研究提出的模型鲁棒性更优。

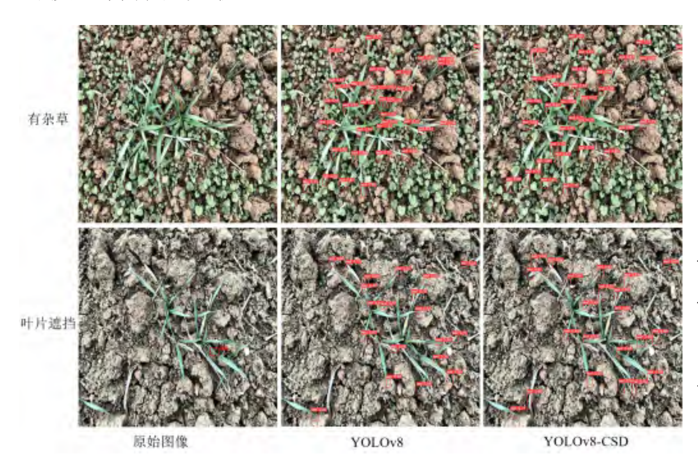


图5 YOLOv8和YOLOv8-CSD对两种复杂场景下小麦叶片图像的检测效果

Fig. 5 The detection effect of YOLOv8 and YOLOv8-CSD on wheat leaf images in two complex scenes

考虑到田间摄像头能够持续捕获时序性的可见光图像，本研究采用其获取图像评估模型改进前后在不同光照条件下对小麦叶尖的检测效果，结果如图6所示。对比YOLOv8模型和YOLOv8-CSD模型对不同光照条件下小麦叶尖检测结果的置信度，YOLOv8-CSD模型检测置信度均有明显提升。

模型改进前后两种典型光照条件下具体检测情况对比结果如表6所示，可知，在晴天环境下YOLOv8模型对小麦叶尖的召回率仅有77%，这是由于在晴天小麦叶片经太阳光照射产生的阴影可能映在真正的叶片上，而模型把叶片阴影当成叶片，导致识别精度和召回率降低。而YOLOv8-CSD模型在两种典型光照条件下对小麦叶尖的召回率均达80%，说明应用本模型能够在不同光照条件下做到自适应调整，模型鲁棒性更好。

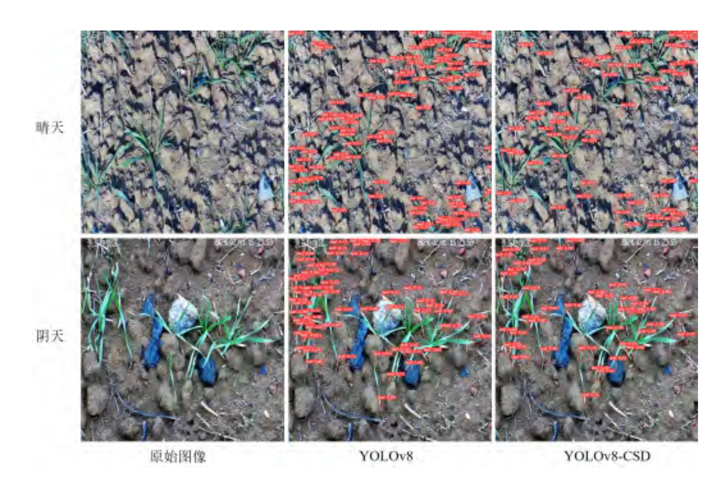


图6 YOLOv8和YOLOv8-CSD对两种典型光照条件下拍摄的小麦叶片图像检测效果图

Fig. 6 Detection effect of wheat leaf images captured under two typical lighting conditions

表6 YOLOv8和YOLOv8-CSD对两种典型光照条件下拍摄的小麦叶片图像检测结果对比

Table 6 Comparison of detection results of wheat leaf images captured under two typical lighting conditions using YOLOv8 and YOLOv8-CSD

光照条件	模型	R/%	$mAP_{0.5}$/%
阴天	YOLOv8	80.6	77.0
	YOLOv8-CSD	82.7	82.2
晴天	YOLOv8	77.0	71.9
	YOLOv8-CSD	80.9	76.1

3.5 模型改进前后不同生长期检测效果对比

鉴于手机拍摄图像的叶尖识别精度高，本研究从测试集中随机选取手机拍摄的小麦在出苗期、分蘖期、越冬期的叶片图像，分别输入YOLOv8m和YOLOv8-CSD中进行实验，观察置信度情况，检测结果如图7所示。其中，YOLOv8-CSD对出苗期和分蘖期小麦叶尖检测框的置信度在0.9左右，相比于YOLOv8模型的0.8有所提升；在越冬期小麦叶片图像的检测中，YOLOv8-CSD检测框的平均置信度为0.8，表现优于标准YOLOv8模型。检测框置信度越高说明模型在学习过程中更好地捕捉到了目标的细节特征，目标存在的可能性越大。本研究提出的YOLOv8-CSD模型嵌入CA注意力机制可以突出重要的通道特征，添加小目标检测层可以提升对细节的敏感度，从而更好地捕捉小麦叶尖。实验结果说明改进后的YOLOv8-CSD相较于原YOLOv8模型具有更好的检测效果，这一结论为后续小麦叶片数检测提供了技术支撑。

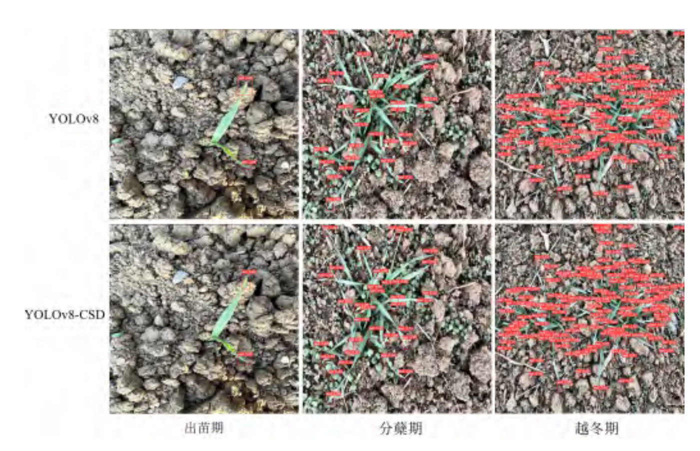


图7　YOLOv8和YOLOv8-CSD对不同生长期小麦叶尖识别效果

Fig. 7 Recognition effects of YOLOv8 and YOLOv8-CSD on the leaf tips of wheat at different growth stages

3.6 复杂大田场景下小麦叶片数检测效果

小麦的叶片数是评估作物生长状况、调节群体结构、指导生产管理的关键指标之一。然而，在实际大田场景下，小麦通常采用无人机撒播，种植密度大，人工识别费时费力且易出现漏检情况。为验证改进后模型的检测效果，选取小麦在出苗期、分蘖期和越冬期三个生长阶段采用1 m双行方式进行叶片数检测拍摄的8张小麦叶片图像，其中出苗期包含103个叶片；分蘖期包含476个叶片；越冬期包含742个叶片。输入训练后的YOLOv8-CSD进行检测，效果如图8所示，可以看出，YOLOv8-CSD模型基本能够有效识别复杂大田场景下小麦叶尖。



a. 出苗期



b. 分蘖期



c. 越冬期

图8　复杂大田场景下YOLOv8-CSD对不同生长期小麦叶片数检测效果

Fig.8 The detection effect of YOLOv8-CSD on the number of wheat leaves with different growth stages in complex field scenes

小麦叶片数具体检测结果统计如表7所示，可知嵌入CA注意力机制的YOLOv8-CSD模型对小麦叶尖的识别性能较好，尤其是在小麦叶片较为稀疏的出苗期，模型能够实现小麦叶尖的有效识别。在分蘖期和越冬期小麦叶片漏检率分别为1.6%和2.8%，这一结果说明本研究所提出的方法泛化能力强，能够满足实际复杂大田场景下小麦叶片数检测的需求。

表7　YOLOv8-CSD对不同生长期小麦叶片数检测结果统计

Table 7 Statistics of the number of wheat leaves detected by YOLOv8-CSD at different growth stages

生长期	检测数/个	漏检数/个	模型检测总时间/s
出苗期	103	0	1.45
分蘖期	468	8	2.27
越冬期	721	21	2.63

4 结　论

针对复杂大田场景下深度学习网络模型在提取特征与学习的能力不足的问题，本研究在YOLOv8模型的基础上进行改进，通过嵌入CA注意力模块来提升模型的感知能力，更好地捕获小麦叶尖目标；修改损失函数，把方向尺度的损失考虑到模型训练中；增加小目标检测层进一步优化特征提取，最终提出了复杂大田场景下的小麦叶尖识别方法。改进后的YOLOv8-CSD模型$mAP_{0.5}$可达85.1%，与主流目标检测模型Faster R-CNN、Mask R-CNN、YOLOv7和SSD相比，模型占用内存最小，但识别精确率最高。相较于原模型，召回率和$mAP_{0.5}$分别提升了6.4%和3.3%。YOLOv8-CSD模型对手机和田间摄像头获取的小麦叶片图像进行检测时，叶尖识别的$mAP_{0.5}$分别达到86.78%和82.43%。研究结果表明，该模型能够有效识别两种不同设备拍摄的小麦叶尖。在复杂大田环境下，YOLOv8-CSD模型具有更好的适应能力，能够有效减少杂草和叶片遮挡情况对叶尖识别的干扰，在不同光照条件下可实现自适应检测，模型鲁棒性强。在叶片稀疏的出苗期，YOLOv8-CSD模型实现了无漏检的检测效果。在叶片密度较大的分蘖期和越冬期，模型依旧保持较低的漏检率，模型泛化性强。

目前，国内外对小麦出苗期、分蘖期以及越冬期的叶片数检测研究较少，本研究通过检测小麦叶尖的数量推断出叶片数，为小麦叶片快速计数提供了可行的方案。所提出的方法能够满足小麦叶尖的有效识别，实现了复杂大田环境下幼苗期小麦叶片数的检测，有助于提高叶片数检测工作的自动化程度，为小麦长势高质量评估提供技术支撑。但改进后模型在对分蘖期以后的小麦叶片检测过程中依旧出现错检和漏检现象，识别准确率仍有提升空间。下一步研究拟补充小麦其他生长期的叶片图像以丰富模型的训练集，提高模型的识别精度和泛化能力，并探索检测模型的轻量化方法以便于边缘端部署。

利益冲突声明： 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] JIN Y C, LIU J Z, XU Z J, et al. Development status and trend of agricultural robot technology[J]. International journal of agricultural and biological engineering, 2021, 14(3): 1-19.

[2] CHEN P P, LI Y, LIU X J, et al. Improving yield prediction based on spatio-temporal deep learning approaches for winter wheat: A case study in Jiangsu province, China[J]. Computers and electronics in agriculture, 2023, 213: ID 108201.

[3] 申雪懿, 李东升, 陈琛, 等. 小麦分蘖数目遗传研究进展与展望[J]. 麦类作物学报, 2023, 43(10): 1344-1350.
SHEN X Y, LI D S, CHEN C, et al. Progress and prospect of genetic research on tiller number in wheat[J]. Journal of triticeae crops, 2023, 43(10): 1344-1350.

[4] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2023: 7464-7475.

[5] ZOU Z X, CHEN K Y, SHI Z W, et al. Object detection in 20 years: A survey[J]. Proceedings of the IEEE, 2023, 111 (3): 257-276.

[6] HE Y, XIAO L. Structured pruning for deep convolutional neural networks: A survey[J]. IEEE transactions on pattern analysis & machine intelligence, 2024, 46(5): 2900-2919.

[7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2016: 770-778.

[8] LU X C, JI J, XING Z Q, et al. Attention and feature fusion SSD for remote sensing object detection[J]. IEEE transactions on instrumentation measurement, 2021, 70: ID 3052575.

[9] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6): 1137-1149.

[10] CHEN W B, LIU M C, ZHAO C J, et al. MTD-YOLO: Multi-task deep convolutional neural network for cherry tomato fruit bunch maturity detection[J]. Computers and electronics in agriculture, 2024, 216: ID 108533..

[11] XU X M, WANG L, LIANG X W, et al. Maize Seedling Leave Counting Based on Semi-Supervised Learning and UAV RGB Images[J]. Sustainability, 2023, 15(12): ID 9583.

[12] DEB M, DHAL K G, DAS A, et al. A CNN-based model to count the leaves of rosette plants (LC-Net)[J]. Scientific reports, 2024, 14: ID 1496.

[13] PRAVEEN KUMAR J, DOMNIC S. Image based leaf segmentation and counting in rosette plants[J]. Information processing in agriculture, 2019, 6(2): 233-246.

[14] YIGIT E, SABANCI K, TOKTAS A, et al. A study on visual features of leaves in plant identification using artificial intelligence techniques[J]. Computers and electronics in agriculture, 2019, 156: 369-377.

[15] XIE X, GE Y, WALIA H, et al. Leaf-counting in monocot plants using deep regression models[J]. Sensors (basel), 2023, 23(4): ID 1890.

[16] KOLHAR S, JAGTAP J. Leaf segmentation and Counting for Phenotyping of Rosette plants using xception-style U-net and Watershed algorithm[M]// Communications in Computer and Information Science. Cham: Springer International Publishing, 2022: 139-150.

[17] REIS D, KUPEC J, HONG J, et al. Real-time flying object detection with YOLOv8[EB/OL]. arXiv: 2305.09972, 2023.

[18] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. Neural netw, 2018, 107: 3-11.

[19] BRAUWERS G, FRASINCAR F. A general survey on attention mechanisms in deep learning[J]. IEEE transactions on knowledge and data engineering, 2023, 35(4): 3279-3298.

[20] ZHU X Z, CHENG D Z, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2019: 6688-6697.

[21] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2021: 13713-13722.

[22] KUNC V, KLÉMA J. Three decades of activations: a comprehensive survey of 400 activation functions for neural networks[EB/OL]. arXiv: 2402.09092, 2024.

[23] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE transactions on cybernetics, 2022, 52(8): 8574-8586.

[24] PENG H, YU S. A systematic IoU-related method: Beyond simplified regression for better localization[J]. IEEE transactions on image processing, 2021, 30: 5032-5044.

[25] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-

CNN[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2017: 2961-2969.
[26] 张阳婷, 黄德启, 王东伟, 等. 基于深度学习的目标检测算法研究与应用综述[J]. 计算机工程与应用, 2023, 59(18): 1-13.
ZHANG Y T, HUANG D Q, WANG D W, et al. Review on research and application of deep learning-based target detection algorithms[J]. Computer engineering and applications, 2023, 59(18): 1-13.

A Rapid Detection Method for Wheat Seedling Leaf Number in Complex Field Scenarios Based on Improved YOLOv8

HOU Yiting[1,2,3], RAO Yuan[1,2,3*], SONG He[4], NIE Zhenjun[1,2,3], WANG Tan[1,2,3], HE Haoxu[1,2,3]

(1. *College of Information and Artificial Intelligence, Anhui Agricultural University, Hefei* 230036, *China*; 2. *Key Laboratory of Agricultural Sensors, Ministry of Agriculture and Rural Affairs, Hefei* 230036, *China*; 3. *Key Laboratory of Smart Agriculture Technology and Equipment in Anhui Province, Hefei* 230036, *China*; 4. *College of Agriculture, Anhui Agricultural University, Hefei* 230036, *China*)

Abstract:

[Objective] The enumeration of wheat leaves is an essential indicator for evaluating the vegetative state of wheat and predicting its yield potential. Currently, the process of wheat leaf counting in field settings is predominantly manual, characterized by being both time-consuming and labor-intensive. Despite advancements, the efficiency and accuracy of existing automated detection and counting methodologies have yet to satisfy the stringent demands of practical agricultural applications. This study aims to develop a method for the rapid quantification of wheat leaves to refine the precision of wheat leaf tip detection.

[Methods] To enhance the accuracy of wheat leaf detection, firstly, an image dataset of wheat leaves across various developmental stages—seedling, tillering, and overwintering—under two distinct lighting conditions and using visible light images sourced from both mobile devices and field camera equipmen, was constructed. Considering the robust feature extraction and multi-scale feature fusion capabilities of YOLOv8 network, the foundational architecture of the proposed model was based on the YOLOv8, to which a coordinate attention mechanism has been integrated. To expedite the model's convergence, the loss functions were optimized. Furthermore, a dedicated small object detection layer was introduced to refine the recognition of wheat leaf tips, which were typically difficult for conventional models to discern due to their small size and resemblance to background elements. This deep learning network was named as YOLOv8-CSD, tailored for the recognition of small targets such as wheat leaf tips, ascertains the leaf count by detecting the number of leaf tips present within the image. A comparative analysis was conducted on the YOLOv8-CSD model in comparison with the original YOLOv8 and six other prominent network architectures, including Faster R-CNN, Mask R-CNN, YOLOv7, and SSD, within a uniform training framework, to evaluate the model's effectiveness. In parallel, the performance of both the original and YOLOv8-CSD models was assessed under challenging conditions, such as the presence of weeds, occlusions, and fluctuating lighting, to emulate complex real-world scenarios. Ultimately, the YOLOv8-CSD model was deployed for wheat leaf number detection in intricate field conditions to confirm its practical applicability and generalization potential.

[Results and Discussions] The research presented a methodology that achieved a recognition precision of 91.6% and an $mAP_{0.5}$ of 85.1% for wheat leaf tips, indicative of its robust detection capabilities. This method exceled in adaptability within complex field environments, featuring an autonomous adjustment mechanism for different lighting conditions, which significantly enhanced the model's robustness. The minimal rate of missed detections in wheat seedlings' leaf counting underscored the method's suitability for wheat leaf tip recognition in intricate field scenarios, consequently elevating the precision of wheat leaf number detection. The sophisticated algorithm embedded within this model had demonstrated a heightened capacity to discern and focus on the unique features of wheat leaf tips during the detection process. This capability was essential for overcoming challenges such as small target sizes, similar background textures, and the intricacies of feature extraction. The model's consistent performance across diverse conditions, including scenarios with weeds, occlusions, and fluctuating lighting, further substantiated its robustness and its readiness for real-world application.

[Conclusions] This research offers a valuable reference for accurately detecting wheat leaf numbers in intricate field conditions, as well as robust technical support for the comprehensive and high-quality assessment of wheat growth.

Key words: wheat leaves; tip detection; leaf counting; attention mechanism; YOLOv8; deep learning

Foundation items: National Natural Science Foundation of China (32371993); Key Research and Development Plan of Anhui Province (202204c06020026, 2023n06020057); Natural Science Research Key Project of Anhui Provincial University (2022AH040125, 2023AH040135)

Biography: HOU Yiting, E-mail: wjx0626fmvp@163.com

*__Corresponding author:__ RAO Yuan, E-mail: raoyuan@ahau.edu.cn